INFORMATION AND COMPUTATION **80**, 249–268 (1989)

Fairness in Context-Free Grammars under Every Choice-Strategy*

SARA PORAT AND NISSIM FRANCEZ

Computer Science Department, Technion–Israel Institute of Technology, Haifa 32000, Israel

In (Porat *et al.*, 1982, *Inform. and Control* **55**, 108–116) the notion of *fair derivations* in context-free grammars was introduced and studied. The main result there is a characterization of fairly terminating grammars as *non-variable-doubling*. In this paper we show that the same characterization is valid under *canonical derivations* in which the next variable to be expanded is deterministically chosen, leaving nondeterminism only to the decision as to which rule (of the chosen variable) to apply. Two families of canonical derivations are introduced and studied as special cases: *spinal* derivations and *layered derivations*.

1. INTRODUCTION

In (Porat *et al.*, 1982) the concept of *fair derivations* in context-free (CF) grammars was introduced in order to study the effects of fairness assumptions in a more abstract context than the usual context of nondeterministic and concurrent programming (Francez, 1986). The main result of that paper is a characterization of fairly terminating CF grammars as *non-variable-doubling* (or *non-expansive*) CF grammars. This characterization establishes the decidability question for fair termination of CF grammars, in contrast to the highly undecidable nature of fair termination in high level nondeterministic programming languages (Harel, 1984).

The motivation for part of the study reported here is based on a dissatisfaction from the way fair behaviors are reflected in that context: fairness can be achieved by applying rules of the same variable in *independent* subderivations. Thus, in tracing the infinite chain of descendants of a specific occurrence of a variable, it need not be the case that indeed all its rules are applied along that chain, as should be the case under some natural conception of "structural fairness" for such derivations. Clearly, all *reduced grammars*, having no useless variables, are fairly terminating under

* A preliminary version of this paper was presented at the STACS 85 conference, Saarbrucken, Germany, LNCS 182 (K. Mehlhom, Ed.), Springer Verlag, 1985.

0890-5401/89 $3.00

the requirement of structural fairness. Thus, the results in Porat *et al.* (1982) mean essentially that structural fairness cannot be reduced to the notion of fairness as introduced in that work. This study began as an attempt towards capturing the desired behavior by considering a different notion of fairness.

CF grammars contain two contexts in which nondeterministic choices are applied in order to determine the next step in a derivation:

(1) The choice of the variable in the sentential form to be replaced.

(2) The choice of the production rule to be applied to the chosen variable.

In this paper we suggest a more restrictive notion of *enabledness* of a production rule by eliminating the first context of nondeterminism: we require to fix deterministically the way the next variable to be replaced is chosen, leaving nondeterminism only in the choice of the next rule to be applied. By this restriction, fewer derivations are considered. The (deterministic) way of chosing the next variable to be replaced is referred to as *choice-strategy*.

A derivation is considered to be *fair under some specific choice-strategy* if for every infinitely often expanded variable, every matching rule of the grammar is applied infinitely often along the derivation. The main result we prove here is that nonvariable doubling is the characteristic property of fair termination under *every* choice-strategy. This obviously means that structural fairness cannot be reduced by fairness under choice-strategy.

For two specific choice-strategies we provide proofs much simpler than for the general case. These special cases represent two specific families of canonical infinite derivations: *spinal derivations* and *layered derivations.*

The results of this paper are comprehensible without prior familiarity with (Porat *et al.*, 1982), though their importance might be better appreciated by readers familiar with the previous treatment.

Another study of fair termination in the context of formal languages, inspired by (Porat *et al.*, 1982), may be found in (Rangarajan and Arunkumar, 1985), where fair termination of EOL systems is studied.

Other abstract models in which the concepts of fairness and fair termination were recently studied are term-rewriting systems (Porat and Francez, 1985), and equational term-rewriting systems (Porat and Francez, 1986). The decidability issue there, in the general case of *ground* term-rewriting systems, is still open.

In Section 2, we define the new notions of fairness and fair termination in the specialized model of context-free grammar that is associated with some choice-strategy. In Section 3 we introduce the techniques of *reconstructions*, as a way of simulating ordinary derivations by derivations under *any* specific choice-strategy. We think that the results presented in this section might be useful in any context dealing with choice-strategies. Some properties that link fair derivations under choice-strategy to their derivation trees are stated in Section 3. The main result is presented in Section 4: fair termination under *every* choice-strategy is characterized in terms of the property of expansiveness. In Section 5, we deal with the two specific families of canonical derivations.

2. Fairness under a Choice-Strategy

In the sequel we use standard notation for CF grammars and languages (Hopcroft and Ullmann, 1979). Let $G=(V, T, P, S)$ be a CF grammar, with no useless variables, i.e.,

(1) $\forall A \in V, \exists \alpha_1, \alpha_2 \in (V \cup T)^*: S \xrightarrow{*} \alpha_1 A \alpha_2$.

(2) $\forall A \in V, \exists w \in T^*: A \xrightarrow{*} w$.

For a given choice-strategy C, we investigate these derivations in which the variable occurrence to be replaced in a derivation step is deterministically chosen under C. These derivations are referred as C-derivations.

In dealing with choice-strategies, all we need is that for every derivation of a sentential form in which there is a variable occurrence, the specific strategy determines exactly one variable occurrence to be replaced in the next derivation step. Following are some possibilities for the form of such strategies.

We can think of a choice-strategy as a terminating deterministic procedure, the input of which is only a sentential form. A well-known strategy studied in the literature is of that kind: for every sentential form, the chosen variable occurrence is the leftmost (or the rightmost). Derivations using that strategy are known to as *leftmost* (or *rightmost*) derivations.

This informal definition of a choice-strategy indicates the properties needed for our theorems. Yet, for the sake of formality, we now give a precise definition of a choice-strategy as a function over grammars and finite derivations.

Definition. A *choice-strategy* C is a function that has two arguments, one is a CF grammar $G=(V, T, P, S)$, and the other is a finite derivation d: $\langle S=\alpha_1 \to \alpha_2 \to \cdots \alpha_n \rangle$ in G. $C(G, d)$ is an ordered pair $\langle A, i \rangle$, where $A \in V$, and there exists some $m>0$ so that $\alpha_n = \beta_1 A \beta_2 A \ldots \beta_m A \beta_{m+1}$ (for $\beta_j \in (V \cup T)^*$, $1 \leqslant j \leqslant m$), and $1 \leqslant i \leqslant m$. We say that the occurrence of A

between β_i and β_{i+1} (the ith *occurrence*) is the *variable occurrence chosen by C*.

A choice-strategy might be one that determines the variable occurrence according to the sentential form and some information about the grammar, like a total ordering on the set of variables. For example: for every sentential form, the chosen variable occurrence is the leftmost occurrence of the "biggest" (in the given order) variable occurrence.

A choice-strategy might depend, in addition to the sentential form, also on the whole (finite) prefix of the derivation from the initial variable up to it. The following choice-strategies are examples to this kind. For the first one, associate with every sentential form a serial number that establishes its position along the derivation. The variable occurrence can be chosen according to the following rule: if the serial number of the form is odd, then the chosen occurrence is the leftmost, otherwise it is the rightmost. For the second strategy, associate with each variable occurrence in a sentential form a natural number, its depth. The chosen variable occurrence is the leftmost from among those unexpanded variable occurrences which have minimal depth.

Note that if the sentential form contains only one variable occurrence, this occurrence is chosen by *every* choice-strategy. A C-derivation starting from a sentential form α and ending with a sentential form β is denoted by $\alpha \xrightarrow{*}_C \beta$.

DEFINITION. (1) (Porat *et al.*, 1982). A production rule $(A \to \alpha) \in P$ is *enabled* in a sentential form β (along a derivation) iff $\beta = \gamma_1 A \gamma_2$, for some $\gamma_1, \gamma_2 \in (V \cup T)^*$.

(2) A production rule $(A \to \alpha) \in P$ is *C-enabled* in a sentential form β (along a C-derivation) iff it is enabled and A is the variable an occurrence of which is chosen as next under the strategy C.

(3) (Porat *et al.*, 1982). A derivation d is *fair* iff it is finite or it is infinite and every rule that is infinitely often enabled along d is also infinitely often applied along d.

(4) A C-derivation d is *C-fair* iff it is finite or it is infinite and every rule that is infinitely often C-enabled along d is also infinitely often applied along d.

(5) (Porat *et al.*, 1982). A CF grammar G is *fairly terminating* iff all its fair derivations are finite.

(6) G is *C-fairly terminating* iff all its C-fair C-derivations are finite.

Remark. For *linear CF grammars* this definition of C-fair C-derivation coincides with the definition of fair derivation as sentential forms in linear grammars contain at most one variable occurrence and have no nondeterminism in variable occurrence choices.

EXAMPLE. Consider the well-known grammar G_1 whose productions are: (1) $S \to aSb$ (2) $S \to \varepsilon$. This grammar is fairly terminating. Indeed, since all its sentential forms are of the form xSy, $x, y \in T^*$, once the rule $S \to \varepsilon$ is applied the derivation terminates. The only infinite derivation is

$$S \xrightarrow{(1)} aSb \xrightarrow{(1)} a^2Sb^2 \xrightarrow{(1)} \cdots \xrightarrow{(1)} a^iSb^i \xrightarrow{(1)} \cdots, \qquad \text{for all } i \geqslant 0,$$

which is clearly unfair. By the above remark G_1 is C-fairly terminating for every choice-strategy C.

EXAMPLE. Let G_2 be given by the following productions:

$$(1) \quad S \to aSSA \qquad (2) \quad S \to \varepsilon \qquad (3) \quad A \to b.$$

Clearly, this grammar is *not* fairly terminating, as is clear from the following infinite fair derivation d_1:

$$S \xrightarrow{(1)} aSSA \xrightarrow{(3)} aSSb \xrightarrow{(2)} aSb \xrightarrow{(1)} aaSSAb$$
$$\xrightarrow{(3)} aaSSbb \xrightarrow{(2)} aaSbb \longrightarrow \cdots$$

All the three rules are applied in a round-robin order, thus d_1 is obviously fair. If the strategy C is that of leftmost derivations, then d_1 is not a C-derivation. Consider the following infinite C-derivation d_2. The chosen variable occurrence, in a sentential form along the derivation, is bold

$$\mathbf{S} \xrightarrow{(1)} a\mathbf{S}SA \xrightarrow{(2)} a\mathbf{S}A \xrightarrow{(1)} aa\mathbf{S}SAA \xrightarrow{(2)} aa\mathbf{S}AA \xrightarrow{(1)} \cdots$$

d_2 is a *C-fair C-derivation*, as in every sentential form along it, the chosen variable occurrence is S, and the two S-rules are infinitely often applied. Thus, for the given choice-strategy C, G_2 is *not* C-fairly terminating. Note that d_2, though being C-fair, is an *unfair derivation*, as the variable A (which is never C-enabled) is infinitely often enabled, but the rule $A \to b$ is never applied. Actually, by the theorem to be proved in the sequel, G_2 is not C-fairly terminating, for *any* choice-strategy C.

3. C-DERIVATIONS AND DERIVATION TREES

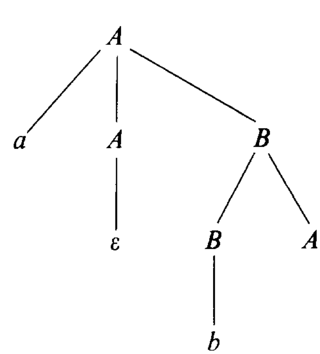
A derivation tree, the root of which is A, consisting of more than one node, may define several possible derivations from A, but the first rule applied along all these derivations is unique. This is the key idea in *reconstructing* a C-derivation using ordinary derivations. The variable occurrence to be replaced in every derivation step along a C-derivation is, of course, chosen according to the choice-strategy C. The rule, to be applied in every derivation step, can be determined according to some given derivation trees. For every reachable sentential form α, $\alpha = x_1 x_2 \cdots x_n$, $x_i \in V \cup T$, a specific derivation tree τ_{x_i} is associated with every occurrence x_i. The idea behind the applied derivation steps, as defined below, is to trace the derivations defined by these trees.

DEFINITION. (1) If d is a partial C-derivation $S \rightarrow_C \cdots \rightarrow_C x_1 x_2 \cdots x_n$, a *continuation* D of d is a finite set of derivation trees $\tau_{x_1}, \tau_{x_2}, \ldots, \tau_{x_n}$ (where x_i is the top symbol of τ_{x_i}).

(2) Let d be a partial C-derivation, D be a continuation for it, and x_i be the variable occurrence chosen by C. An elementary *step of C-reconstruction* is a transition $\langle d, D \rangle \rightarrow_C \langle d', D' \rangle$, where d' is the extension of d by applying the first rule defined by τ_{x_i}, and D' is the residual continuation.

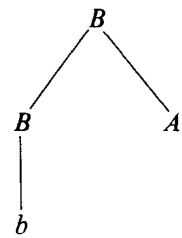
(3) The *C-reconstructed derivation* determined by $\langle d, D \rangle$ is the limit of the d_i in the maximal sequence $\langle d, D \rangle = \langle d_0, D_0 \rangle \rightarrow_C \langle d_1, D_1 \rangle \rightarrow_C \cdots$.

EXAMPLE. Let $\alpha = AaAabBc$ be the last sentential form in a partial C-derivation d, and let the continuation D be the set $\{\tau_A^1, \tau_a^2, \tau_A^3, \tau_a^4, \tau_b^5, \tau_B^6, \tau_c^7\}$. Assume that the second occurrence of A in α is the variable occurrence chosen by C, and let τ_A^3 be the derivation tree



After one step of C-reconstruction, the derived sentential form is $AaaABabBc$ and the associated continuation is: $\{\tau_A^1, \tau_a^2, \tau_a^{3,1}, \tau_A^{3,2}, \tau_B^{3,3}, \tau_a^4, \tau_b^5, \tau_B^6, \tau_c^7\}$, where $\tau_a^{3,1}$ consists of the single node a, $\tau_A^{3,2}$ is

and $\tau_A^{3,3}$ is



Consider the derivation tree τ_d of an infinite C-fair C-derivation d. Note that τ_d defines several possible derivations in G, but only one C-derivation, the one that can be C-reconstructed from $\langle S, \tau_d \rangle$.

The following remarks capture the relation between the infinite C-fair C-derivation d, and the tree τ_d.

Remark 1. The production rule $A \to \alpha$ is C-enabled along d iff there is an internal node in τ_d labeled by A. This is so due to the fact that if an A-production rule is C-enabled, then there is a derivation step along d in which A is the variable an occurrence of which is chosen as next under C, thus this occurrence is expanded. (Note that if there is a leaf in τ_d labeled by A, then the production rule $A \to \alpha$ is *enabled* along d.)

Remark 2. The production rule $A \to \alpha$ is C-enabled k times (or infinitely many times) along d iff there are k (or infinitely many) distinct internal nodes in τ_d, labeled by A.

Remark 3. The production rule $A \to x_1 x_2 \cdots x_n$ is applied k times (or infinitely many times) along d iff there are k (or infinitely many) distinct internal nodes in τ_d, all labeled by A, and the successors of which are exactly $x_1, x_2, \ldots, x_n$ in the same ordering imposed by the rule. In such case we say that the rule $A \to x_1 x_2 \cdots x_n$ *occurs* in the tree k times (or infinitely many times).

Conclusion. If the variable A labels infinitely many internal nodes in τ_d, then, by the fairness assumption of d, every A-rule occurs infinitely many times in τ_d.

The following remarks present the relation between τ_d, and some derivations in G defined by this tree.

643/80/3-5

Remark 4. If v_1 is an internal node in τ_d labeled by A_1, and v_2 is a node labeled by A_2 in the subtree, the root of which is v_1, then there are $\alpha_1, \alpha_2 \in (V \cup T)^*$, so that $A_1 \xrightarrow{*} \alpha_1 A_2 \alpha_2$.

Remark 5. Let v_1 be an internal node in τ_d labeled by A_1. Let u_1 and u_2 be two nodes, labeled by A_2 and A_3, in the subtree the root of which is v_1, so that for $i = 1, 2$, u_i is not a node in the subtree the root of which is $u_{(i \bmod 2)+1}$. Then, there are $\alpha_1, \alpha_2, \alpha_3 \in (V \cup T)^*$, such that $A_1 \xrightarrow{*} \alpha_1 A_2 \alpha_2 A_3 \alpha_3$ or $A_1 \xrightarrow{*} \alpha_1 A_3 \alpha_2 A_2 \alpha_3$.

We say that a rule $A \to x_1 x_2 \cdots x_n$ occurs along some path π in a derivation tree, if the rule occurs in the tree, such that the occurrence of the l.h.s. A is on π (thus, there is also some x_i that occurs on π).

4. Variable-Doubling and *C*-Fair-Termination

Definition. A CF grammar is *variable doubling* (*expansive*) iff there is a variable $A \in V$ such that $A \xrightarrow{*} \alpha_1 A \alpha_2 A \alpha_3$ for some $\alpha_1, \alpha_2, \alpha_3 \in (V \cup T)^*$.

Example. Consider the following grammar G_3 (Table I). This grammar is variable-doubling (A-doubling), as is seen from the derivation tree presented in Fig. 1.

The main theorem in (Porat *et al.*, 1982) is the following: a CF grammar is fairly terminating iff it is nonvariable-doubling. The main result in the sequel is the proof that for *every* C, nonvariable-doubling is a necessary and sufficient condition for C-fair termination. This result cannot be obtained directly from the original proof.

Establishing the (only-if) direction of the equivalence proof in (Porat *et al.*, 1982) is quite simple. If a grammar is expansive, say A-doubling for some variable $A \in V$, then one occurrence of A can be used for the doubling process, whereas the other one can take care of fairly applying all the B-rules, for every variable B that can be derived from A. Thus, we obtain

TABLE I

The Variable-Doubling Grammar G_3

$$G_3 = (\{S, A, B, D\}, \{a, b, c\}, P, S)$$

where P contains the rules:

$$S \to A$$
$$A \to BSD \mid b$$
$$B \to aA$$
$$D \to c$$

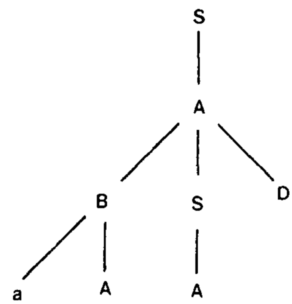


FIG. 1. Derivation tree of A-doubling in G_3.

an infinite fair derivation. This construction is strongly dependent on the ability to expand *every* variable occurrence in a sentential form. Thus, for a given choice-strategy C, the constructed derivation is not necessarily a C-derivation. The other direction of the proof is derived by some stronger results, involving *derivation forests* and *induced forests*. A close look at this proof shows that it relies on the fact that every subtree of a derivation tree defines a derivation. This is not the case for subtrees of a C-derivation tree.

Note that in contrast to the situation in the case of arbitrary derivations, in spite of the fact that the grammar is A-doubling, there need not exist a sentential form C-derivable from A in which A occurs twice. Moreover, deriving a sentential form α with two occurrences of A, still cannot ensure both of them to be expanded in a C-derivation starting from α. Our main concern in the sequel is to overcome this obstacle by reconstructions. We relate to several finite derivations trees.

DEFINITION. For a given CF grammar G:

(1) An $A-B$ *reachability* tree for $A, B \in V$, is a finite A-rooted derivation tree with B as the only nonterminal leaf symbol. Note that if $A = B$ an $A-A$ reachability tree may consist of only one node.

(2) An *A-doubling* tree, for $A \in V$, is a finite A-rooted derivation tree with exactly two nonterminal leaf symbols, both of them being A. (A variable A is doubling itself iff there is an A-doubling tree.)

(3) An *A-terminal* tree, for $A \in V$, is a finite A-rooted derivation tree with no nonterminal leaf symbols.

THEOREM (Necessity of nonvariable-doubling for C-fair termination). *For* every *CF grammar G, and for* every *choice strategy C: If G is C-fairly terminating, then G is not variable-doubling.*

Proof. We have to show that if the grammar G has a variable doubling itself, then for every choice-strategy C, there is an infinite C-fair C-derivation. The proof is constructive and is based on C-reconstruction using finite derivation trees.

Let $A \in V$ be a variable that doubles itself. Let $V' = \{A_1, A_2, ..., A_n\} \subseteq V$ be the set of variables that can be derived in G from A. In other words,

$$V' = \{B \in V \mid \exists \alpha_1, \alpha_2 \in (V \cup T)^*: A \overset{*}{\rightarrow} \alpha_1 B \alpha_2\}.$$

The set V' is not empty, since at least $A \in V'$. We partition the set V' to two disjoint sets V'_1 and V'_2.

$$V'_1 = \{B \in V' \mid \exists \alpha_1, \alpha_2 \in (V \cup T)^*: B \overset{*}{\rightarrow} \alpha_1 A \alpha_2\}, \qquad V'_2 = V - V'_1.$$

The set V'_1 is not empty, since at least $A \in V'_1$. The set V'_2 may be empty.

The constructed infinite C-derivation, as described in the sequel, assures the cyclic application of all of those rules, the l.h.s. of which is some variable in V', infinitely many times. This is done by C-reconstruction using the following finite derivation trees as elements in the continuations: an $S-A$ reachability tree, an $A-B$ reachability tree, a $B-A$ reachability tree, and a B-doubling tree, for every $B \in V'_1$ (the existence of the required doubling trees follows from the definition of V' and V'_1), and a B-terminal tree, for every $B \in V'$.

We now impose the ordering in which the rules are to be applied as a round robin, ensuring fairness. For every $B \in V'$, let $\{B \rightarrow \alpha_1^B, B \rightarrow \alpha_2^B, ..., B \rightarrow \alpha_{n_B}^B\}$ be all the B-rules, enumerated in some arbitrary, but fixed, ordering.

The construction of the infinite C-fair C-derivation is done in two stages:

Stage 1. Ensuring fairness with respect to those rules, the l.h.s. of which is some variable in V'_1. Let $A_1, ..., A_k$ be all the variables in V'_1. We use two counters l and m.

Step 1. We associate with S as its continuation an $S-A$ reachability tree, and apply steps of C-reconstruction. The C-reconstructed derivation ends with a sentential form in which the variable occurrence to be replaced is the leaf A of the $S-A$ reachability tree.

Step 2. $l \leftarrow 0$ (initializing the counter for the variables in V'_1; starting a cycle).

Step 3. $l \leftarrow l+1$. We update the continuation s.t. the tree associated with the chosen occurrence of A is an $A-A_l$ reachability tree, and apply steps of C-reconstruction. The C-reconstructed derivation now ends with a sentential form in which the variable occurrence to be replaced is the leaf A_l in the $A-A_l$ reachability tree.

$m \leftarrow 0$ (initializing the counter for the A_l-rules).

Step 4. We update the continuation s.t. the tree associated with the chosen occurrence A_l is an A_l-doubling tree. After some finite number of steps of C-reconstruction the variable occurrence to be replaced is an occurrence of A_l, that is, a leaf in the A_l-doubling tree.

$$m \leftarrow m+1.$$

We associate with the chosen occurrence of A_l, an A_l-terminal tree that defines derivations in G in which the first rule applied along them is $A_l \rightarrow \alpha_m^{A_l}$. A step of C-reconstruction ensures the application of this A_l-rule. After some finite number of such steps, the variable occurrence to be replaced is again A_l—the second leaf in the given A_l-doubling tree.

Step 5. If $m < n_{A_l}$, go to step 4.

Step 6. All the A_l-rules were already applied (in this cycle). We now associate with the chosen occurrence A_l, an $A_l - A$ reachability tree. After some finite number of steps of C-reconstruction, the variable occurrence to be replaced is the leaf A of the $A_l - A$ reachability tree.

Step 7. If $l < k$, go to step 3. Otherwise, each rule, the l.h.s. of which is some variable in V_1', has been applied at least once by executing step 4 (thus completing a cycle).

The required C-derivation consists of an indefinite repetition of steps (2) up to (7).

Stage 2. Ensuring fairness with respect to those rules, the l.h.s. of which is some variable in V_2'. Let τ be the infinite tree C-reconstructed in Stage 1. Consider an internal node in τ labeled by $B \in V_2'$. By the definitions of V_1' and V_2', for every $B' \in V_1'$, B' cannot be derived from B in G. Thus, the subtree, the root of which is the considered internal node, is a finite tree, and if we replace it by some B-terminal tree, we still have an infinite tree that defines a C-derivation that is C-fair with respect to the variables in V_1'.

For every $B \in V_2'$, we use a counter l^B for the B-rules. At the beginning, all these counters are initialized to zero. Starting from the pair $\langle S, \tau \rangle$, one stops the C-reconstruction each time some $B \in V_2'$ is selected for expansion, and then applies the following steps:

(1) $l^B \leftarrow l^B + 1$.

(2) Change the continuation s.t. the tree associated with the chosen occurrence B is a B-terminal tree that defines derivations in G, in which the first rule applied along them is $B \rightarrow \alpha_{l^B}^B$ (thus, ensuring the application of this B-rule).

(3) If $l^B = n_B$ (all the B-rules have been applied the same number of times), then $l^B \leftarrow 0$ (initializing again the counter in order to obtain repetitive application of all the B-rules).

It is easy to prove that every rule enabled infinitely often along the constructed C-derivation, is really applied infinitely often, thus we obtain the required infinite C-fair C-derivation.

THEOREM (Sufficiency of nonvariable-doubling for C-fair termination). *For* every *CF grammar* G, *and for* every *choice-strategy* C: *If* G *is not variable-doubling, then* G *is* C*-fairly terminating.*

Proof. We show that if there exists an infinite C-fair C-derivation d, then there is some variable that can double itself.

DEFINITION. A node v is a *dominating root* if there is an infinite path in the tree, starting from v, with infinitely many occurrences of the variable that labels v, and no node v', such that there is a path from v' to v, satisfies the condition.

In order to clarify this definition and the following complicated proof, let us consider the grammar G_4 (Table II). Let C be the strategy discussed above, for which a serial number is associated with every sentential form. This number establishes the form's position along the derivation. The variable occurrence to be replaced is the leftmost for odd forms, and the rightmost for even ones. One can easily prove that the derivation tree in Fig. 2 corresponds to an infinite C-fair C-derivation. In this tree both successors of the root are dominating roots. We use this as a running example throughout the proof.

CLAIM 1. *In every infinite derivation tree, there is only a finite number of dominating roots.*

TABLE II

The Grammar G_4

$$G_4 = (\{S, A, B, D\}, \{a, b\}, P, S)$$

where P contains the rules:

$$S \to AB$$
$$A \to BAB \mid a$$
$$B \to ABA \mid bD$$
$$D \to b$$

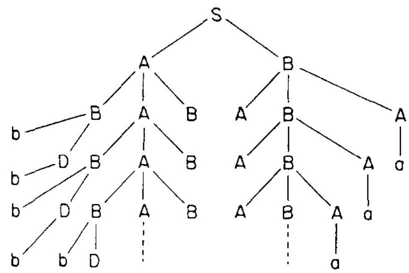


FIG. 2. Derivation tree of an infinite C-fair C-derivation in G_4.

Proof. If a node v is a dominating root, then no node that is on the path from the root to v is also a dominating root. Thus, consider the tree τ' obtained from the original one τ by retaining only arcs of nodes that are not dominating roots. In other words, τ is truncated at all nodes that are dominating roots. Assume, by way of contradiction, that there are infinitely many dominating roots, thus τ' is an infinite tree. As τ' is finitely branching, then by Konig's lemma, there is an infinite path the nodes of which are all not dominating roots. As the set of variables (labeling nodes on the tree) is finite, we get a contradiction to the definition of a dominating root.

CLAIM 2. *There is only a finite number of nodes that do not belong to subtrees, the roots of which are dominating roots.*

Proof. By the finiteness of the truncated tree defined in the proof of Claim 1.

With every dominating root v, labeled by A, we associate an infinite path π_v, that starts from v, and contains infinitely many occurrences of A (its existence is established by the definition of a dominating root). We refer to this path as the *dominating path* of the subtree, the root of which is v.

In the derivation tree of Fig. 2, for every dominating root there is only one infinite path that starts from it, thus this path is defined to be the dominating path of the subtree.

We now define a partition of V:

$$V = V^1 \cup \cdots \cup V^n, \qquad \text{where} \quad A \in V^i \text{ iff } i = \min\{j \mid A \xrightarrow{j} w \in T^*\}.$$

Clearly, as there are no useless variables, every variable is covered by this partition. In G_4, for example, $V^1 = \{A, D\}$; $V^2 = \{B\}$; $V^3 = \{S\}$.

For every dominating root v, we denote by A_v, one of the variables B, such that $B \in V^i$, B occurs infinitely often along the dominating path π_v, and for every other variable B', $B' \in V^j$, occurring infinitely often along π_v, $i \leqslant j$.

CLAIM 3. *For every dominating root v, there is some A_v-rule that occurs only finitely often along the dominating path π_v. We refer to this A_v-rule as the* decreasing rule *of A_v.*

Proof. If $A_v \in V^1$, then the claim immediately follows. A_v has a rule with a terminal r.h.s., and such a rule can never occur along an infinite path.

If $A_v \in V^i$, $i > 1$, then by the definition of the partition, there is some derivation of length i, that starts from A_v, and ends with some terminal word. Let $A_v \rightarrow \alpha$ be the first production rule applied along this derivation. Obviously, for every variable B that occurs in α, there is some derivation of length less than i, that starts from B and ends with some terminal word, thus $B \in V^j$, where $j < i$. Suppose, by way of contradiction, that this A_v-rule occurs infinitely often along π_v. Thus, there is some variable B, $B \in V^j, j < i$, that occurs infinitely often along π_v, contradicting the minimality in the definition of A_v.

In the tree of Fig. 2, let v_1, v_2 be the two dominating roots, labeled by A, B, respectively. A_{v_1} is the variable A, A_{v_2} is B, and $A \rightarrow a$, $B \rightarrow bD$ are, respectively, their decreasing rules.

By the above conclusion, based on the C-fairness assumption of the given C-derivation d, for every dominating root v, as A_v labels infinitely many nodes along the dominating path π_v, the decreasing rule of A_v occurs infinitely many times in τ_d. By the above claims (1 and 2), there is some dominating root v', such that the decreasing rule of A_v occurs infinitely many times in the subtree, the root of which is v'.

We now define a binary relation R on dominating roots. For the two dominating roots v and v', vRv' iff the decreasing rule of A_v occurs infinitely many times in the subtree, the root of which is v'.

Since there is only a finite number of dominating roots, there are $l \geqslant 1$ dominating roots: $v_1, v_2, \ldots, v_l$, so that $v_i R v_{(i \bmod l)+1}$, for every $1 \leqslant i \leqslant l$. We refer to this set of dominating roots as a *cyclic dominating set* of the tree. In the tree of Fig. 2, both dominating roots are in a cyclic dominating set of the tree. For a cyclic dominating set $v_1, \ldots, v_l$, let A_i be the variable labeling the dominating root v_i, $1 \leqslant i \leqslant l$.

CLAIM 4. *There is some i, $1 \leqslant i \leqslant l$, such that the decreasing rule of A_{v_i} occurs only finitely often along the dominating path $\pi_{v_{(i \bmod l)+1}}$ (though infinitely many times in the subtree, the root of which is $v_{(i \bmod l)+1}$).*

Proof. If there is some i, $1 \leqslant i \leqslant l$, so that $A_{v_i} \in V^1$, then the claim immediately follows (as the decreasing rule of this A_{v_i} does not occur along any dominating path, which is infinite by definition).

Otherwise, assume by way of contradiction, that for every i, $1 \leqslant i \leqslant l$, the decreasing rule of A_{v_i} occurs infinitely often along the dominating path $\pi_{v_{(i \bmod l)+1}}$. Let $A_{v_i} \in V^{j_i}$, where $j_i > 1$. By the assumption, for every i, there is some variable $B_i \in V^{k_i}$, having an occurrence on the r.h.s. of the decreasing rule of A_{v_i}, thus $k_i < j_i$ and B_i labels infinitely many nodes along the dominating path $\pi_{v_{(i \bmod l)+1}}$. By the definition of $A_{v_{(i \bmod l)+1}}$, we get that $j_{(i \bmod l)+1} \leqslant k_i$, and so $j_{(i \bmod l)+1} < j_i$. By the cyclicity, we get the contradiction $j_{(i \bmod l)+1} < j_i$ and $j_i < j_{(i \bmod l)+1}$.

Claim 5. *Consider the dominating root v_i, such that the decreasing rule of A_{v_i} occurs only finitely often along the dominating path $\pi_{v_{(i \bmod l)+1}}$ (its existence is established by the last claim). The variable $A_{(i \bmod l)+1}$ doubles itself.*

Proof. By the definition of a dominating path, Claim 4 and Remark 5, there are $\alpha_1, \alpha_2, \alpha_3 \in (V \cup T)^*$, such that

$$A_{(i \bmod l)+1} \overset{*}{\to} \alpha_1 A_{(i \bmod l)+1} \alpha_2 A_{v_i} \alpha_3 \quad \text{or} \quad A_{(i \bmod l)+1} \overset{*}{\to} \alpha_1 A_{v_i} \alpha_2 A_{(i \bmod l)+1} \alpha_3.$$

By the definition of A_v and Remark 4, for every i, $1 \leqslant i \leqslant l$, there are some β_1^i, β_2^i, so that $A_{v_i} \overset{*}{\to} \beta_1^i A_i \beta_2^i$. By the definition of R and Remark 4, for every i, $1 \leqslant i \leqslant l$, there are some γ_1^i, γ_2^i, such that $A_{(i \bmod l)+1} \overset{*}{\to} \gamma_1^i A_i \gamma_2^i$. By the cyclicity of the dominating set, we can easily conclude, that for every i, j, $1 \leqslant i$, $j \leqslant l$, there are some $\delta_1^{i,j}$, $\delta_2^{i,j} \in (V \cup T)^*$, such that $A_{v_i} \overset{*}{\to} \delta_1^{i,j} A_j \delta_2^{i,j}$. So, there is a sentential form, derivable from A_{v_i}, in which there is an occurrence of $A_{(i \bmod l)+1}$.

The last claim completes the proof of the theorem.

5. Families of Canonical Derivations

In this section we deal with two specific families of canonical infinite derivations: spinal derivations and layered derivations. For each family we consider a representative choice-strategy C, such that the C-derivations all belong to the corresponding family.

Spinal derivations are (infinite) derivations in the derivation trees of which there is only one infinite path. Known examples of such derivatings are the *leftmost* and the *rightmost* derivations. In these examples, the descendant variable occurrences of any given variable occurrence in a form are replaced before any "sibling" occurrence is replaced. As a representative of this family we shall consider the leftmost derivations.

DEFINITION. A variable occurrence A is *next* (chosen) under the leftmost strategy (L) in a sentential form β iff $\beta = wA\gamma$ for $w \in T^*$, $\gamma \in (V \cup T)^*$.

We now use L instead of the generic C in the definitions above. According to this definition of a strategy, a rule is L-enabled on a sentential form whenever its l.h.s. variable has an occurrence that is the leftmost variable in the form.

EXAMPLE. We present an example of an infinite L-unfair L-derivation in G_3 (defined in Table I):

$$S \xrightarrow[L]{} A \xrightarrow[L]{} BSD \xrightarrow[L]{} aASD \xrightarrow[L]{} aBSDSD \xrightarrow[L]{*} a^i A(SD)^i$$

$$\xrightarrow[L]{} a^i B(SD)^{i+1} \cdots \qquad \text{for} \quad i \geqslant 2.$$

This infinite L-derivation is L-unfair since A is infinitely often the next (to be replaced), but the rule $A \to b$ is *never* applied.

As a consequence of the above theorems we have: For every CF grammar G, G is L-fairly terminating iff G is not variable doubling. This characterization of L-fair termination, can be proved in a way simpler than the general one.

The (If) Direction. The general form of a derivation tree of an infinite L-derivation is as shown in Fig. 3. It contains exactly one infinite path referred to as the *spine*. The subtrees to the left of the spine are all finite. To

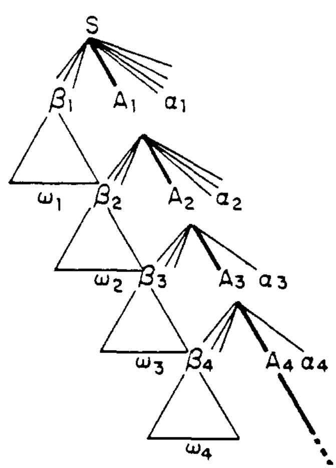


FIG. 3. An infinite L-derivation tree.

the right of the spine are only leaves labeled either by terminals or by variables. Using the above terminology, every derivation tree of an infinite L-fair L-derivation has only one dominating root v, labeled by some variable $A \in V$. The spine defines the dominating path π_v. By Claim 4, the decreasing rule of A_v occurs only finitely often along the dominating path π_v. Thus, A_v occurs infinitely many times in finite subtrees to the left of the spine. As A_v can derive A, we get that the variable A doubles itself.

The (Only If) Direction. Let $A \in V$, $\alpha_1, \alpha_2, \alpha_3 \in (V \cup T)^*$ such that $A \xrightarrow{*} \alpha_1 A \alpha_2 A \alpha_3$. The idea is to construct an infinite L-fair L-derivation in such a way that A is "responsible" for the fair application of rules. We design the derivation so that A is encountered alternately; once in a left finite subtree

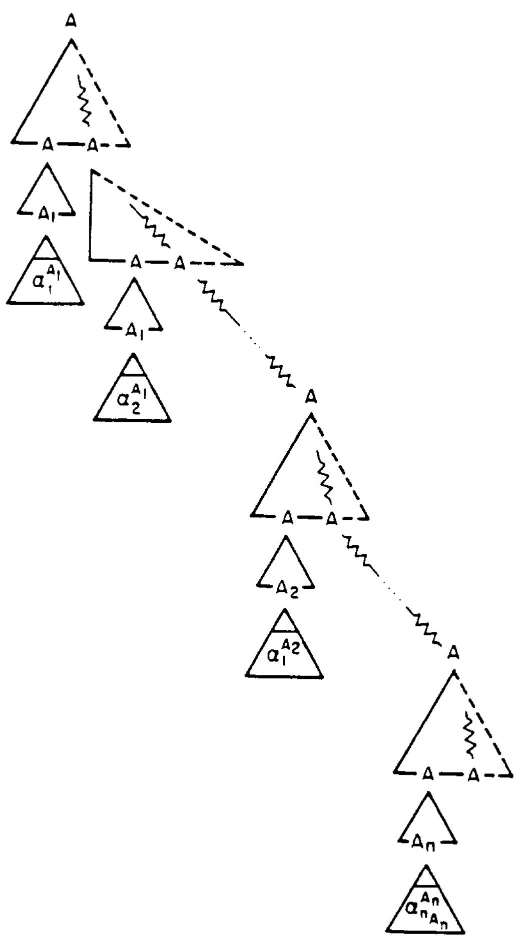


FIG. 4. A full cycle.

and the second time down the spine. This is possible due to the A-doubling property.

Using, again, the above terminology, we construct a full-cycle ensuring application of all the B-rules, for $B \in V'$. This is done by reconstructing an A-doubling tree, an $A-B$ reachability tree, and some B-terminal trees, for every $B \in V'$. The construction of the infinite L-fair L-derivation is simpler than in the general case, as the choice-strategy L imposes a fix ordering between the two leaves of the doubling tree of A. A full-cycle is demonstrated in Fig. 4.

By prefixing to an indefinite repetition of a full cycle an L-derivation reconstructing an $S-A$ reachability tree, we obtain the required infinite L-fair L-derivation.

Layered derivations are (infinite) derivations in the derivation trees of which the leaves are always labeled by terminals. The variable occurrences are replaced in such a way that no variable occurrence is left unexpanded for ever. We consider, as a representative of this family, derivations where replacements are performed in an order dictated by the depth; for variable occurrences that are in the same depth a left-to-right order is imposed.

To express formally this strategy denoted by LA, we associate with each variable occurrence in a sentential form a natural number, its depth.

DEFINITION. A variable occurrence A is *next* in a form α under the strategy LA iff $\beta = \gamma A \delta$, $\gamma, \delta \in (V \cup T)^*$, and there exists a natural number i such that the depth of the occurrence of A to the right of γ is i, and the depths of all the variable occurrences in γ are $i+1$ and these in δ are i.

The general form of an LA-derivation tree is shown in Fig. 5. In the figure, a small square denotes a terminal-labeled node while a small circle

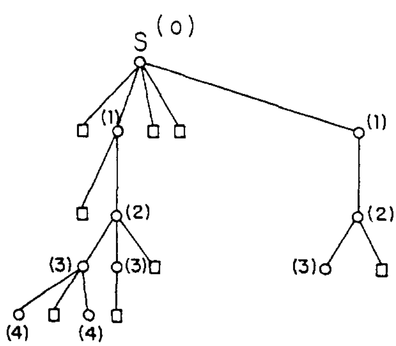


FIG. 5. An LA-derivation tree.

denotes a variable-labeled node. The depths of the nodes are also indicated. Again, as a consequence of the above theorems, we have:

> For every CF grammar G, G is LA-fairly-terminating iff G is not variable-doubling.

This conclusion can again be proved in a way simpler than the general one.

The (If) Direction. This direction is immediate following from the characterization theorem in (Porat *et al.*, 1982), as every infinite LA-fair LA-derivation is also fair under the definition there.

The (Only If) Direction. Suppose the given grammar doubles the variable A. We first describe a section of an LA-derivation starting with A and guaranteeing that for any variable B derivable from A all the B-rules are used. This can be done since the variable-doubling property ensures two occurrences of A (though not necessarily at the same layer); the left one is used for fairly expanding all the variable occurrences while the right one allows another expansion of A. The details of the formal construction are similar to the spinal case and all one has to check is that they can be carried out in using LA-derivations. We omit the details.

As before, this section is repeated infinitely often prefixed with another section generating a sentential form containing a (first) A. Here one has to take care that the form-portions appearing to the left of A and to its right are expanded in such a way as to produce finite subtrees (with terminal leaves). This is possible by the assumption of the absence of useless variables.

ACKNOWLEDGMENT

The part of the second author was partially supported by the fund for the promotion of research at the Technion, Israel Institute of Technology.

RECEIVED June 30, 1986; ACCEPTED January 22, 1988

REFERENCES

FRANCEZ, N. (1986), "Fairness," Springer-Verlag, New York/Berlin.

HAREL, D. (1984), A general result about infinite trees and its applications, *in* "Proceedings, 16th ACM-STOC, Washington, D.C., May."

HOPCROFT, J. E., AND ULLMANN, J. D. (1979), "Introduction to Automata Theory, Languages and Computation," Addison–Wesley, Reading, MA.

PORAT, S., AND FRANCEZ, N. (1985), Fairness in term-rewriting systems, *in* "Rewriting Techniques and Applications, Dijon, France, May"; Lecture Notes in Computer Science Vol. 202 (J. P. Jouannoud, Ed.), pp. 287–300, Springer-Verlag, New York/Berlin.

PORAT, S., AND FRANCEZ, N. (1986), Full-communication and fair-termination in equational (and combined) term-rewriting systems, *in* "CADE-8, Oxford, England, July."

PORAT, S., FRANCEZ, N., MORAN, S., AND ZAKS, S. (1982), Fair derivations in context free grammars, *Inform. and Control* **55**, 108–116.

RANGARAJAN, K., AND ARUN-KUMAR, S. (1985), Fair derivations in EOL systems, *Inform. Process. Lett.* **20**, 183–188.